IDA PAPER P-2098

PASSIVE RANGING

Irvin W. Kay

August 1988

Prepared for
Office of the Under Secretary of Defense for Acquisition

INSTITUTE FOR DEFENSE ANALYSES
1801 N. Beauregard Street, Alexandria, Virginia 22311-1772

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY N/A	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE N/A	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) IDA Paper P-2098	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Institute for Defense Analyses 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION DoD-IDA Management Office, OUSD(A)
6c. ADDRESS (City, State, and Zip Code) 1801 N. Beauregard Street Alexandria, VA 22311	7b. ADDRESS (CITY, STATE, AND ZIP CODE) 1801 N. Beauregard Street Alexandria, VA 22311
8a. NAME OF FUNDING/SPONSORING ORGANIZATION DUSD (R&AT) 8b. OFFICE SYMBOL (If applicable) EST	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER MDA 903 84 C 0031
8c. ADDRESS (City, State, and Zip Code) The Pentagon, Room 3D359 Washington, DC 20301	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT / PROJECT NO. / TASK NO. T-D2-210 / WORK UNIT ACCESSION NO.
11. TITLE (Include Security Classification) Passive Ranging	
12. PERSONAL AUTHOR(S). Irvin W. Kay	
13. TYPE OF REPORT Final 13b. TIME COVERED FROM 9/87 TO 8/88 14. DATE OF REPORT (Year, Month, Day) August 1988 15. PAGE COUNT 47	
16. SUPPLEMENTARY NOTATION	

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	range finding, passive systems, optical images, infrared images, dynamic images, moving sensor, airborne, optical flow, photogrammetry, errors

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

This paper compares the classical photogrammetric approach to the passive ranging problem with the approach based on the theory of optical flow applied to dynamic imagery. It also describes various errors, due to the environment and imperfections of the sensor system, long recognized by photogrammetrists but largely ignored by proponents of the optical flow methodology.

An error sensitivity analysis indicates that the optical flow approach is more sensitive to errors--and to those due to image distortion, in particular--than is the photogrammetric approach. This is primarily because of the necessarily small line-of-sight angle between two views of the same object point by the sensor, especially when the sensor platform is constrained to move nearly parallel to the optical axis of the system.

Although this constraint on the platform motion guarantees that the focus of expansion (FOE) remains in the field of view, that fact does not make the system less sensitive to error, but rather more so. Analysis shows that the opposite constraint imposed by photogrammetrists, having the platform move in a direction orthogonal to the optical axis, reduces the error sensitivity even though it forces the FOE out of the field of view--in fact, to infinity.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL Jeffrey F. Nicoll	22b. TELEPHONE (Include Area Code) (703) 578-2987 22c. OFFICE SYMBOL

IDA PAPER P-2098

PASSIVE RANGING

Irvin W. Kay

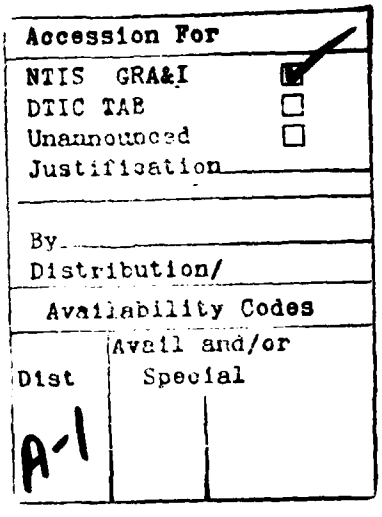
August 1988

INSTITUTE FOR DEFENSE ANALYSES

Contract MDA 903 84 C 0031
Task T-D2-210

FOREWORD

This paper was prepared as part of Task T-D2-210 under the technical cognizance of Dr. John M. MacCallum, Office of the Deputy Under Secretary of Defense (Research and Advanced Technology/Electronic Systems Technology).



ABSTRACT

This paper compares the classical photogrammetric approach to the passive ranging problem with the approach based on the theory of optical flow applied to dynamic imagery. It also describes various errors, due to the environment and imperfections of the sensor system, long recognized by photogrammetrists but largely ignored by proponents of the optical flow methodology.

An error sensitivity analysis indicates that the optical flow approach is more sensitive to errors--and to those due to image distortion, in particular--than is the photogrammetric approach. This is primarily because of the necessarily small line-of-sight angle between two views of the same object point by the sensor, especially when the sensor platform is constrained to move nearly parallel to the optical axis of the system.

Although this constraint on the platform motion guarantees that the focus of expansion (FOE) remains in the field of view, that fact does not make the system less sensitive to error, but rather more so. Analysis shows that the opposite constraint imposed by photogrammetrists, having the platform move in a direction orthogonal to the optical axis, reduces the error sensitivity even though it forces the FOE out of the field of view--in fact, to infinity.

CONTENTS

I. INTRODUCTION

This paper analyzes the problem of estimating the range of a remote object by means of image data that a passive sensor acquires in a moving airborne platform. The objective is to characterize the nature of additional information needed for the range calculation and to examine the effect of optical system errors on its accuracy.

The science of photogrammetry, for which the preferred sensor has been the aerial photographic camera, has dealt with similar problems for more than a century. Although aimed at an application served by a different type of sensor, the present analysis is entirely consistent with the methods of classical photogrammetry, the theory and practice of which Ref. 1 describes in considerable detail.

The application of interest here is real time passive ranging, for which the forward-looking infrared (FLIR) device is a more appropriate sensor than the aerial camera. Reference 2 proposes a computer-driven system to solve the ranging problem by means of algorithms that exploit dynamic imagery acquired continuously during the platform motion, rather than the static data recorded on film by traditional photogrammetric equipment.

The approach of Ref. 2 is one developed and refined in recent years by a research community interested in diverse areas of science and modern technology, such as the nature of human visual perception and various engineering applications involving pattern recognition and automatic scene analysis.

The approach has two operational phases. The goal of the first phase is to solve a pattern recognition problem: the computer must select some number of points, each of which has identifiable images in at least two different frames of the changing scene. With the aid of this data and certain other information, the goal of the second phase is to derive for a particular point a range estimate relative to the instantaneous sensor position associated with one of the frames.

The second-phase calculation relies on certain geometrical information that the methodology can infer from the way the image data changes as the platform and sensor move. Investigators call the continuous image change "optical flow," the properties of

which are quite simple when the sensor motion is strictly translational. In fact, optical flow in the absence of rotation behaves according to the rules of perspective (cf. Ref. 3) known to artists for hundreds of years.

Obviously, both phases of the passive ranging operation contribute errors to the range estimate. However, proponents of the dynamic imagery methodology (cf. Ref. 2), when concerned about errors at all, have concentrated on those arising in the first phase, ignoring for the most part those associated with the second.

Fortunately, classical photogrammetrists, not having to contend with the first phase, have paid considerably more attention to errors ascribed to equipment and the environment that affect the second phase. Reference 1 deals at length with this subject.

A primary objective of this paper is to relate the fund of knowledge built up over the years by photogrammetric experience to the problems still to be faced in applications of the dynamic imagery methodology. As a preliminary step toward this goal, Section II reviews some relevant topics in basic optical theory and photogrammetry.

Section III presents an independent analysis of the optical flow phenomenon in the context of classical geometrical optics, obtaining results equivalent to some given in the literature concerned with dynamic imagery methodology. It also discusses implications of the results for applications such as passive ranging and compares them with similar implications that have influenced the standard methods of photogrammetry.

Section IV lists the error sources discussed in Ref. 1, which also describes instrument calibration techniques for correcting them. Section V analyzes the propagation of image parameter errors due in part to such sources and discusses their influence on the accuracy of range estimates obtained by means of optical flow methods. The error propagation analysis of Section V results in Eqs. (22) and (23), which give the fractional error in the range estimate explicitly in terms of all image parameter errors, including those derived from the effects noted in Section IV.

Section VI, using results from Section V, derives the effect of optical distortion on the error in estimating the range of an object point on the ground observed at two different instants by a sensor moving at constant velocity parallel to the ground. For the case treated in Section VI, it is assumed that the optical axis of the sensor, the velocity vector, and the vertical are coplanar and that the object point also lies in the same plane. Appendix B generalizes this treatment to include the case in which the object point does not lie in the plane determined by the optical axis and the velocity vector.

Section VII briefly summarizes some earlier conclusions, including the minimum data requirements for passive ranging and the limitations of image analysis for that purpose. It also recalls some observations made in this document about error sensitivity and some conclusions about what steps are necessary in acquiring image data for accurate range estimation.

Appendix A derives a method for calculating the instantaneous translation and rotation of the sensor in terms of optical flow data. Unlike others found in the literature, this method, which appears to be new, has the advantage of providing the results in closed form.

II. CLASSICAL OPTICS AND PHOTOGRAMMETRY

Passive ranging is the operation of estimating the distance between an object in space and a local reference point relative to a passive imaging sensor that supplies the requisite data. The algorithms used for the calculation must be consistent with geometrical optics, which provides the classical mathematical model for an optical instrument or imaging system.

According to that theory, the ideal imaging system generates a one-to-one mapping of the points in a three-dimensional object space to the points in a three-dimensional image space by means of rays that obey the laws of refraction and reflection. Ideally, each object point defines a pencil of rays which, when traced through the system, eventually converge to form another pencil centered at the corresponding image point.

Geometrical optics, which determines the rays connecting object points with image points, predicts that, except for certain image points called stigmatic images, not all of the rays in a ray bundle will intersect a single point. Thus, in general, the ray bundles associated with image points are only approximately pencils. However, for well-corrected optical instruments and for image points on a particular surface, ideally a plane, the deviation of the actual ray bundles from ideal pencils in image space is physically negligible.

In fact, the first-order paraxial approximation, the limit in which the angle that every ray makes with the optical axis approaches zero, also known as Gaussian optics, depicts the performance of a well-designed imaging sensor accurately enough for most purposes. Nevertheless, certain errors that do occur are important in the passive ranging application. If not corrected in the range calculations, these errors, to be discussed later, should be taken into account in assessing the limitations of the process.

It is assumed that the optical system is cylindrically symmetric about an optical axis and that the propagation medium in the image space is the same as that in the object space. Then, according to Gaussian optics, the instrument is equivalent to the geometrical construct illustrated in Fig. 1.

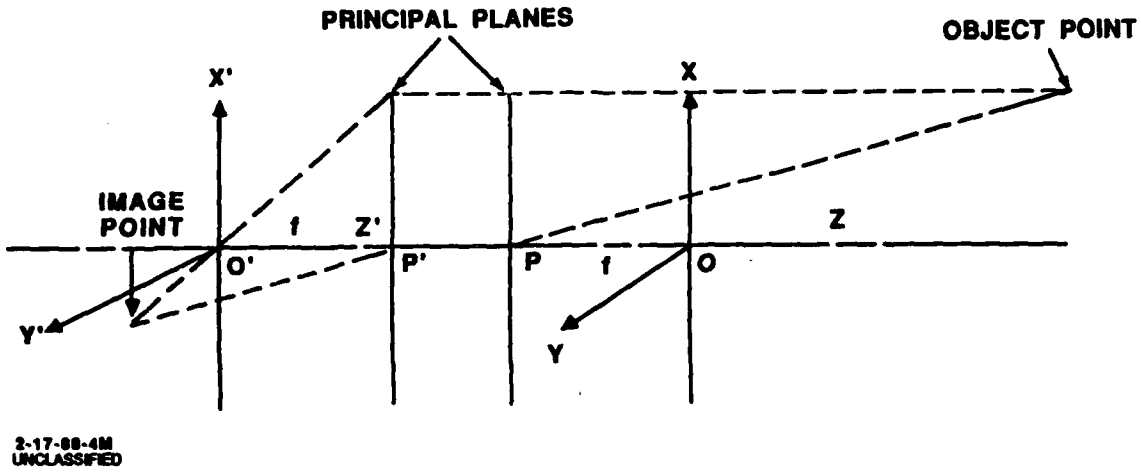


Figure 1. Optical Coordinate Systems.

The front and back principal planes associated with the instrument are perpendicular to the optical axis Z and intersect it at the front and back nodal points P and P'. The front and back focal planes are also perpendicular to Z and intersect it at O and O'. Because of the stated assumptions, the system has a single focal length f equal to the distance between O and P, which in turn is equal to the distance between O' and P'.

It is customary to use separate (right-handed) Cartesian coordinate systems to define points in the image and object spaces. The object-space Z axis and the image-space Z' axis are both coincident with the optical axis. The object-space X-Y plane is the front focal plane at O, and the image-space X'-Y' plane is the back focal plane at O'. The X' axis is parallel to the X axis, and the Y' axis is parallel to the Y axis.

According to Newton's lens law, if Z is the Z coordinate of an object point and Z' is the Z' coordinate of the corresponding image point, then one has

$$ZZ' = -f^2 , \tag{1}$$

which defines the conjugate planes where corresponding object and image points lie. Also, according to the first-order optical theory, the principal ray connecting an object point with P maps into a parallel ray connecting P' with the corresponding image point in image space.

If the object point moves to infinity, its position is defined by the direction of a corresponding principal ray through P. The conjugate image plane becomes the back focal

plane, and the principal ray, which maps into a parallel ray through P' in image space, determines the image point in the focal plane.

The projective transformation of every object point into an image point on a conjugate image plane by a sensor's optical system provides the data that is the basis for classical photogrammetry. From the image-plane coordinates of the same image point viewed by two sensors in different locations, it is theoretically possible to determine the object point's object-space coordinates relative to either sensor if the transformation from the coordinate systems associated with one sensor to those associated with the other is known. Photogrammetrists refer to two such images of the same point as a stereoscopic pair.

In fact, if several points occur simultaneously in a stereoscopic pair of scenes, it is also theoretically possible to reconstruct the required coordinate transformation from the combined image data and some additional information. This additional information usually includes the distances between certain object-space control points.

The basis for locating an object point by means of two images is geometrically simple. The process can be described in terms of the first-order optics model of an imaging sensor as follows.

When the sensor is focused on a region of points in object space, the front nodal point P is the center for the projection of the object points onto the conjugate image plane, the distance of which from the back principal plane is presumed known. It is customary to replace the true image plane conceptually by an equivalent one that is the same distance on the other side of the front principal plane. Then the image of any object point occurs where the ray connecting the front nodal point to the object point intersects the equivalent image plane.

If the same object point is viewed from two different sensor positions, the two projection rays that determine the corresponding image points associated with the sensor positions must both intersect that object point. Since each ray is determined by a single point image relative to one sensor coordinate system and the rays can intersect in only one point, the two images must determine the location of the object point uniquely. Figure 2 illustrates the geometrical configuration underlying this argument.

The most general rigid-body coordinate transformation from one sensor position to the other consists of a translation, defined by three parameters, and a rotation, given by an orthogonal matrix, which is also defined by three parameters. Thus, the transformation

from the coordinate system associated with one sensor to that associated with the other is specified by six parameters.

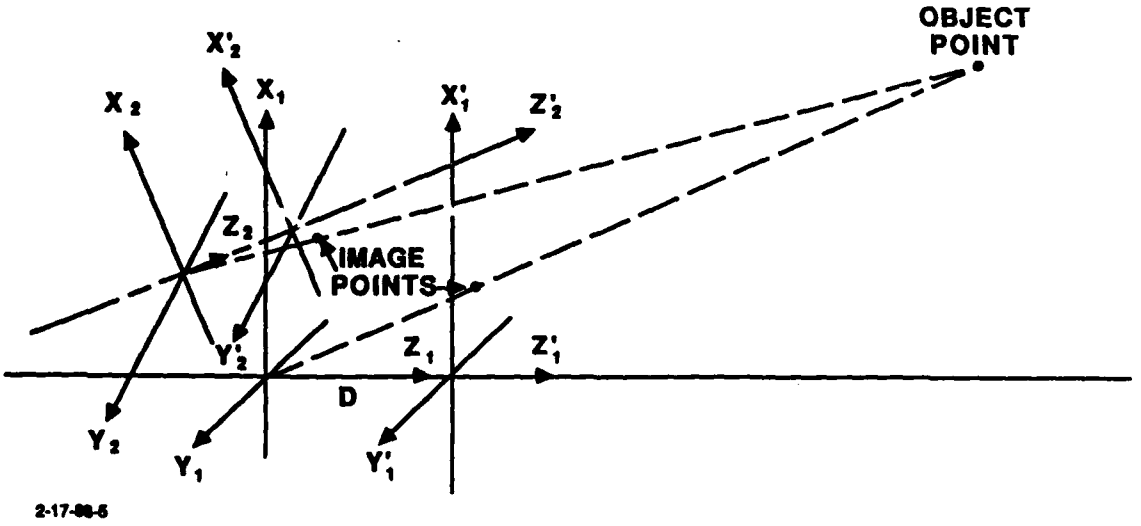


Figure 2. Locating an Object Point by Means of Two Image Points.

A standard photogrammetric procedure for determining the transformation between the coordinate systems associated with a stereoscopic pair depends on the ray projections used to locate an object point and is based on a general relation, called the coplanarity equation (cf. Ref. 1, p. 55 ff.), implied by the geometric constraints inherent in the projections. The equation is an analytical statement of the fact that the three lines connecting the object point and the front nodal points that are the centers of projection for the two sensors form a triangle and must therefore lie in the same plane. Every pair of rays corresponding to a pair of images of some object point viewed at both sensor positions determines a coplanarity equation in this way.

That is, if $\mathbf{R}_1$ is the coordinate vector from the nodal point to the object point when the system is in position 1, $\mathbf{R}_2$ is the corresponding object point coordinate vector when the optical system is in position 2, and $\mathbf{T}$ is the translation vector from the position 1 nodal point to the position 2 nodal point, then the coplanarity condition is

$$\mathbf{T} \cdot \mathbf{R}_1 \times \mathbf{R}_2 = 0.$$

The equations derived from coplanarity are nonlinear, and all are homogeneous in at least one unknown. Thus, in principle, five or more such ray pairs should suffice to determine five of the six parameters that characterize the coordinate transformation, i.e., up to a scale factor. The photogrammetric practice is to use the coplanarity relation to construct a two-dimensional image map of an object-space region but to use other procedures to derive the map scale.

Another geometrical fact, that the nodal, image, and object points on a ray must be collinear, leads to the relation called the collinearity equation (cf. Ref. 1, p. 88 ff.), which is used as an alternative to the coplanarity equation for some purposes. In particular, e.g., in the Church method of space resection (Ref. 1, p. 60 ff.), it is customary to determine the scale of a map by means of the collinearity relation applied to three known object-space control points.

Equation (6) in Section III can be regarded as equivalent to the collinearity condition in the limit of a small difference between the two coordinate systems associated with a stereoscopic pair. However, some of the quantities in the corresponding system of equations in Ref. 1 are expressed in object-space coordinates rather than the image-space coordinates used in Eq. (6).

Reference 1 describes several numerically efficient methods for solving the nonlinear systems of equations resulting from the coplanarity and collinearity conditions, as well as statistical methods for dealing with random errors and noisy data. It is geometrically evident that the estimated object-point position given by a pair of rays projected from the corresponding image points at two different sensor positions is less sensitive to error if the angle between the rays intersecting at the object point is wide rather than narrow. This is more likely to be the case if the angular coordinates of the two corresponding image points relative to the projection center and optical axis at each sensor position are markedly different rather than nearly the same.

III. OPTICAL FLOW METHODS

From a point of view somewhat different from that of classical photogrammetry, this section addresses the problem of determining the coordinate transformation between the two image frames of a stereoscopic pair. The purpose, however, is still the same: to derive sufficient relations between the image data and the parameters defining the coordinate transformation to make it possible to deduce the transformation from the image data. Consistent with the photogrammetric experience cited in Section II, it will be found that only five of the parameters necessary to characterize the rotation and translation of the coordinate system can be obtained in this way. The magnitude of the translation vector will then remain to be determined by other means.

In recent years, an interest in the optical theory of nontraditional disciplines such as image analysis and pattern recognition, with such nontraditional applications as the modeling of human or animal vision and robotics, has fostered another approach to goals that are similar to if not identical with those of photogrammetry. The basis for this approach is a phenomenon called optical flow, which refers to the apparent motion of points in the image plane of a sensor that is moving relative to object space.

In particular, if the sensor motion consists entirely of a linear rigid-body translation, the image points appear to flow away from or toward a single point called the focus of expansion (FOE), which remains stationary. The FOE is clearly the point where an extension of the sensor velocity vector at the nodal point would intersect the image plane.

When the sensor also undergoes a change of attitude due to a rotation about some axis, the resulting optical flow in general will obscure the true FOE associated with the translational motion alone. For the purpose of deriving the correct FOE, methods for extracting the optical data supplied by the translational component of the motion have appeared in a number of papers in the last decade.

As just observed, a knowledge of the FOE is equivalent to a knowledge of the translation direction, i.e., it supplies only two of the six parameters needed to locate the position of an object point by means of data derived from two different scenes that contain

its image. When this is the objective, proponents of optical flow methodology must supplement the calculation with additional information at their disposal or derived from other geometrical properties of the viewed images.

It is obvious that the optical flow approach is valid in principle. Any pair of image-plane displays obtained at two different instants while a sensor moves will satisfy the basic photogrammetric requirement for two different views of the same object point. Therefore, as observed earlier, if five or more recognizable points occur in both scenes, the standard methods of analytical photogrammetry will yield most but not all of the data needed to determine completely the most general rigid-body motion of an imaging system. In fact, the image data supplied by a sensor in continuous motion over time is similar to data derived from a sequence of static photographs and used in the aerotriangulation process (cf. Ref. 1, Chapter IX) for accurately scaled terrain mapping by means of aerial photography.

Deriving the relationship between optical flow and the rigid-body motion of an imaging system is a straightforward matter. For this purpose it is convenient to use a coordinate sysem for which the X-Y plane is coincident with the front principal plane instead of the object-space coordinate system described earlier and shown in Fig. 1. It is also convenient to represent geometric position and motion in terms of vectors: with the Cartesian coordinates (x, y, z) of a point given by the components of a radius vector $x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$, where $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ are unit vectors along the X, Y, and Z coordinate axes.

The front nodal point, which is also the origin of the coordinate system, will be regarded here as the center of the optical system's motion. The instantaneous velocity $\dot{\mathbf{R}}$ of any point that is fixed in space is given by

$$\dot{\mathbf{R}} = -\mathbf{v} + \mathbf{R} \times \boldsymbol{\omega} \quad , \tag{2}$$

where $\mathbf{v}$ is the translational velocity of the origin and $\boldsymbol{\omega}$ is the rotational velocity of the coordinate system about some axis through the origin.

To avoid confusion between the image-plane distance and the focal length, as indicated in Fig. 2, the distance of the (equivalent) image plane from the origin will be designated by D here, although it has become customary to use f for this purpose. This precaution is mainly cosmetic, since for most cases of interest replacing D by the focal length, for which f is also the usual symbol, would create a negligible error in a range estimate.

Lower case letters will be used to designate image-point radius vectors. That is, a vector r with components (u,v) in the image plane is understood to be the vector with components (u,v,d).

From the projective construction of the image point corresponding to an object point, it is geometrically evident that the image of an object point given by the vector

$$\mathbf{R} = X\mathbf{i} + Y\mathbf{j} + Z\mathbf{k} \tag{3}$$

is given by the vector

$$\mathbf{r} = D\mathbf{R}/Z\,. \tag{4}$$

From (4) it follows that

$$\dot{\mathbf{r}} = D(Z\dot{\mathbf{R}} - \dot{Z}\mathbf{R})/Z^2 = -D\mathbf{k} \times (\mathbf{R} \times \dot{\mathbf{R}})/Z^2\,. \tag{5}$$

On substituting from (5) into (2) and applying standard vector identities, it follows that

$$\dot{\mathbf{r}} = D[-Z\mathbf{v} + v_z\,\mathbf{R} - (\mathbf{R} \cdot \boldsymbol{\omega})(\mathbf{k} \times \mathbf{R}) + R^2\,(\mathbf{k} \times \boldsymbol{\omega})]/Z^2\,.$$

After substituting from (4), this becomes

$$\dot{\mathbf{r}} = -D\mathbf{v}/Z - v_z\mathbf{r}/Z + (\mathbf{r} \cdot \boldsymbol{\omega})(\mathbf{k} \times \mathbf{r})/D + r^2\,(\mathbf{k} \times \boldsymbol{\omega})/D\,. \tag{6}$$

The vector differential equation (6) defines the optical flow of all image points at an instant of time in terms of the instantaneous translation and rotation velocities of the sensor. It is equivalent to a pair of scalar differential equations for the image-point coordinates:

$$\begin{aligned} \dot{x} &= -Dv_x/Z + xv_z/Z + xy\omega_x/D - (x^2 + D^2)\,\omega_y/D + y\omega_z\,, \\ \dot{y} &= -Dv_y/Z - yv_z/Z + (y^2 + D^2)\,\omega_x/D - xy\omega_y/D - x\omega_z\,. \end{aligned} \tag{7}$$

Since the FOE, by definition, is the point that remains stationary when the sensor does not rotate, its location can be obtained from (7) by setting the left-hand sides and the components of $\boldsymbol{\omega}$ equal to zero. That is, the FOE will have the image coordinates $(Dv_x/v_z, Dv_y/v_z)$. In fact, with zero rotation and nonzero translation, for any image point the optical

flow line obtained by extension from the instantaneous flow direction at the point will have the equation, in terms of (x_1,y_1) coordinates measured in units equal to D:

$$y_1 = v_y/v_z + m(x_1 - v_x/v_z) \ , \tag{8}$$

where one has

$$m = \frac{dy}{dx} \ ,$$

evaluated at the image point. All such lines obtained in the same way from all image points will intersect at the FOE.

In principle, in the general case when rotation does occur, (7) provides the means of determining some but not all of the unknown velocity and rotation vector components. In practice, it would be necessary to solve the differential equations numerically in terms of image measurements made over a time interval short enough to insure that the instantaneous velocities do not change significantly. Otherwise, the FOE and the velocities would vary during the interval and would have to be treated as functions of time.

Thus, (7) can be regarded as equivalent to an algebraic system of equations obtained by multiplying both sides by the time interval dt. Then the derivatives on the left-hand side can be replaced by distances dx and dy. Similarly, on the right-hand side the velocity vector components can be replaced by displacement components, and the rotation vector components can be replaced by angular displacements.

Although it should be kept in mind that the unknown velocity vectors have been replaced by displacements, with little danger of confusion the notation for their components can remain the same. Also, it is evident, after dividing both sides of the equations by D, that D can be set equal to 1; this amounts to an understanding that image distances are to be expressed in units equal to D.

It is possible to eliminate the unknown distance z from the two equations by transposing all terms on the right-hand side containing rotation components to the left-hand side and dividing the second equation by the first. The result is

$$\frac{dy - (y^2 + 1)\,\omega_x + xy\omega_y + x\omega_z}{dx - xy\omega_x + (x^2 + 1)\,\omega_y - y\omega_z} = \frac{v_y - yv_z}{v_x - xv_z} \ . \tag{9}$$

For every mutual image point in the two frames observed by the sensor, it is possible to write an equation of the form given by (9). On dividing the numerator and

denominator of the right-hand side by the magnitude v of v it is evident that, whatever the number of the image points contributing, the resulting system of equations can only determine the direction of v, and not its magnitude. Thus, as in the case of the photogrammetric equations based on coplanarity, five unknowns associated with the sensor's rotation and its translation direction can be determined, in principle, by data obtained from a minimum of five image points.

Dividing by v_z instead of v and multiplying through by the denominators on both sides reduces the system to a system of quadratic equations for the unknowns:

$$(y^2+1)v_1\omega_x - xyv_1\omega_y - xv_1\omega_z - xyv_2\omega_x + (x^2+1)\,v_2\omega_y - yv_2\omega_z - x\omega_x - y\omega_y$$
$$+ (x^2+y^2)\,\omega_z - v_1dy + v_2dx + xdy - ydx = 0\,, \qquad (10)$$

where it has been assumed that v_z is not zero, v_1 has been written for the ratio v_x/v_z, and v_2 has been written for the ratio v_y/v_z. Geometrically, a solution corresponds to the intersection of five four-dimensional quadric manifolds, which, in fact, have hyperbolic two-dimensional cross sections. Thus, with the use of only five image points for data the solution may not be unique.

If as many as 11 image points contribute to the data in (10), thereby furnishing a system of 11 equations, the problem changes from solving a quadratic system to solving one that is linear. It is only necessary to treat as independent unknowns the quantities $v_1\omega_x$, $v_1\omega_y$, $v_1\omega_z$, $v_2\omega_x$, $v_2\omega_y$, $v_2\omega_x$, ω_x, ω_y, ω_z, v_1, v_2.

Reference 4 uses relations that are essentially the same as (7), along with the fact that the linear extensions of the displacements (dx, dy) at all image points must intersect at the FOE,[1] to separate the translational and rotational effects of the sensor motion on the optical flow. The calculation involves adjusting the components of ω to minimize an error function defined, for a set of extended lines in the flow directions, as the mean square separation of the points where the lines in the set intersect a single extended flow line chosen arbitrarily. This process determines the FOE coordinates if the set contains enough lines.

The method depends on a sequence of iterations which, with luck, will converge reasonably quickly to a result that gives the true FOE position. Although some numerical examples encourage the hope that this will be the case in general, confidence in the method

---

1 This fact is not independent of (7). It is used only to define an error function that the solution algorithm of Ref. 4 attempts to minimize.

requires faith not only in its quick convergence but also in its convergence to the correct result. That the calculation provides the true FOE is not a foregone conclusion because, although a minimum for the error function clearly must exist, it is conceivable that the function may have two or more local minima corresponding to points other than the FOE.

Knowing the FOE position is equivalent to knowing the direction but not the magnitude of the coordinate-system translation during the sensor motion. The solution of (10), or the calculation of Ref. 4, also provides an estimate of the coordinate-system rotation, the effect of which can be removed from the combined coordinate transformation. However, Appendix A of this paper derives what appears to be a more efficacious method for calculating the coordinate-system translation direction and rotation than the approaches suggested thus far.

In any event, whatever the means employed, the first objective is to acquire the coordinate-system transformation and remove the rotational part. Then, for a given object point, if $\mathbf{r}$ and $\mathbf{r}'$ are radius vectors to the corresponding image points in the frames before and after the translation, the standard photogrammetric procedure described in Section II can be used as follows to locate the object point.

The condition that lines in the directions of $\mathbf{r}$ and $\mathbf{r}'$ must intersect at the object point can be expressed as

$$\mathbf{R} = \mathbf{v} + \mathbf{R}' \ , \tag{11}$$

where $\mathbf{R}$ and $\mathbf{R}'$ are vectors to the object point from the projection centers before and after the translation and where $\mathbf{v}$ is the translation vector, which is the product of the translational velocity and the time interval of the motion. The vectors $\mathbf{R}$ and $\mathbf{R}'$ have the same directions as the image-point vectors $\mathbf{r}$ and $\mathbf{r}'$.

The objective is to calculate the range, which is the magnitude R' of the vector $\mathbf{R}'$. For this purpose, the following pair of equations can be obtained from (11), first by taking the cross product of both sides with $\mathbf{v}$ and then by taking the cross product of both sides with $\mathbf{r}'_0$, which is the unit vector in the direction of $\mathbf{R}'$:

$$\begin{aligned} \mathbf{R}\mathbf{v} \times \mathbf{r}_0 &= \mathbf{R}'\mathbf{v} \times \mathbf{r}'_0 \ , \\ \mathbf{R}\mathbf{r}_0 \times \mathbf{r}'_0 &= \mathbf{v} \times \mathbf{r}'_0 \ . \end{aligned} \tag{12}$$

Relative to some polar coordinate system in the plane defined by the three vectors $\mathbf{R}$, $\mathbf{R}'$, $\mathbf{v}$, the respective angular coordinates of the vectors will be denoted by θ, θ', θ_v. Then, since all of the cross products in (12) are mutually parallel vectors (orthogonal to the

plane determined by $\mathbf{R}$, $\mathbf{R}'$, $\mathbf{v}$), the equations in (12) are equivalent to the two scalar equations

$$R\sin(\theta_v - \theta) = R'\sin(\theta_v - \theta') \ ,$$

$$R\sin(\theta - \theta') = v\sin(\theta_v - \theta') \ . \tag{13}$$

Substituting from the second equation of (13) into the first and solving leads to the formula

$$R' = \frac{v\sin(\theta_v - \theta)}{\sin(\theta - \theta')} \tag{14}$$

for the range.

Except for notation, Eq. (14) is identical to the formula for range given in Ref. 2, which proposes to use algorithms based on optical flow to do passive ranging with a FLIR from an aircraft platform. It is clear from (14) that the magnitude and the direction of the platform's translational velocity are required to accomplish this end. However, as observed earlier, optical flow data can only yield the velocity direction. Therefore, additional information, such as the velocity magnitude or some kind of ground truth in terms of known object control points like those in ordinary photogrammetric usage, is also necessary.

IV. ERRORS

As remarked at the end of Section II, when the observation of overlapping image data takes place in scenes viewed from directions that are nearly the same, object-point location estimates derived from the data will be particularly sensitive to measurement errors. This condition is inherent in the optical flow methodology because it relies on data due to small changes in image-point positions. Therefore, measurement errors are, if anything, more important for a passive ranging technique based on optical flow than for one based on classical photogrammetry, the standard procedures of which are aimed at insuring the opposite condition.

Reference 1 (e.g., pp. 478-483, 486-488) deals extensively with error contributions to photogrammetry. However, the literature concerned with similar applications of image measurement, using other approaches such as optical flow theory, has largely ignored such considerations.

Optical systems inevitably contribute errors that are important in these applications, for passive ranging in particular. As indicated in Ref. 1, it is not only necessary to correct certain errors that the sensor causes, but in order to do so it is also necessary to obtain the data for that purpose by applying specific calibration procedures to the optical equipment. Reference 1 (pp. 232-274) treats in some detail standard methods of calibrating aerial cameras.

Reference 1 lists the following errors, most of which must occur in any optical system, no matter what its purpose may be. Some appear specific to the particular equipment used in photogrammetry; however, these also have their counterparts in other imaging devices.

A. PRINCIPAL POINT COORDINATE DISPLACEMENTS

Errors in both the focal length and the optical center, referred to as the principal point, are significant and occur even in aerial cameras providing fiduciary marks that supposedly point toward the center. In addition, image distortion may require an

adjustment from what is called the effective focal length of the optical system to a more operationally accurate value called the calibrated focal length.

A preflight calibration can determine these effects quantitatively. The resulting data must be applied to adjusting coordinates relative to the coordinate systems used in processing the image data observed by the sensor.

B. DISTORTION

Image distortion results primarily from two sources: imperfect lens design and decentered lens elements. It is also customary to distinguish two types: radial, which is symmetric about the optical center in the image plane, and tangential, which is only symmetric about a line through the optical center.

Analytical distortion models, involving a small number of parameters that can be obtained from calibration measurements, are useful both for correcting image data and for estimating the effect of distortion on image errors. For the radial type, the model is usually a polynomial in odd powers of the radial distance from the optical center. For the tangential type, the model involves a polynomial in both odd and even powers of the radial distance multiplied by a linear combination of the sine and cosine of the angle measured from a line emanating from the optical center.

C. EMULSION DEFORMATION

This error source occurs with aerial cameras because photographic material, film or glass, is subject to deformation between exposure and development. Although a sensor, such as a FLIR, that records or transmits image data by other than photographic means will not suffer from precisely the same effect, it will undoubtedly encounter one that is similar enough to require the same kind of correction. For example, errors may exist in the alignment or lateral positioning of detector elements.

With a scanning device nonuniformities in the scan motion may also produce an image distortion effect reminiscent of that due to film emulsion. To correct distortion resulting from dynamic sources, it may be necessary to make in-flight calibration measurements for the requisite data.

D. PLATEN CURVATURE

The camera pressure platen that is supposed to keep the film in a plane orthogonal to the optical axis may be imperfectly machined and therefore have small undulations. An array of detectors in an IR sensor may have a similar defect.

E. ATMOSPHERIC REFRACTION

Variations in the atmospheric index of refraction cause rays from object points to bend in accordance with Snell's law. Generally, the index of refraction is horizontally stratified; i.e., it varies with height rather than laterally. Aside from a static index of refraction distribution that persists for relatively long periods of time, there are also rapid fluctuations, due to turbulence, that produce statistical fluctuations in the ray pathlengths. Not only can this effect cause defocusing, but it may also shift a focus, i.e., an image point, laterally.

V. ERROR PROPAGATION

When a function depends upon two or more independent parameters that have small errors, the corresponding error in the function is approximately equal to the sum of the errors contributed by the individual parameters. To the same order of approximation, the error contributed by a single parameter is equal to the parameter error multiplied by the derivative of the function with respect to the parameter. To obtain the fractional (or percent) error, it is only necessary to replace the derivatives by logarithmic derivatives.

Thus, the fractional error ε in the range estimate derived from optical flow theory and given by (14) is approximately equal to the sum of the error contributions e_v, $e_{\theta\theta'}$, $e_{v\theta}$ from v and the angle differences $\theta - \theta'$, $\theta_v - \theta$. That is

$$\varepsilon = e_v + e_{v\theta} + e_{\theta\theta'} \ , \tag{15}$$

where

$$e_v = \varepsilon_v/v, \ e_{v\theta} = \varepsilon_{v\theta} \cot(\theta_v - \theta), \ e_{\theta\theta'} = -\varepsilon_{\theta\theta'} \cot(\theta - \theta') \ ,$$

and the quantities ε_v, $\varepsilon_{v\theta}$, $\varepsilon_{\theta\theta'}$ are the absolute errors in v and the angle differences $\theta_v - \theta$, $\theta - \theta'$.

It is obvious from (15) that when the angle differences are small, the fractional range error is very sensitive to errors in those differences. On the other hand, in the first frame the angle between the optical axis and the ray to the object point will be small when the field of view is narrow, e.g., in the passive ranging case. These factors seem to contradict a remark in Ref. 2 to the effect that the algorithm used in the passive ranging method employed there is optimal when the angle between the velocity direction and the optical axis is near zero.

At the very least, the term "optimal" appears to be an exaggeration. As indicated in Ref. 1, for the ideal photogrammetric condition the view is vertical (p. 53) and the platform motion is horizontal (p. 55), which is a rule of thumb that (15) supports.

That is, if the velocity direction is perpendicular rather than parallel to the optical axis, then the angular difference $\theta_v - \theta$ will be close to 90 deg. Thus, in this case on the

right-hand side of (15), the second term, which is the only one that depends explicitly on the velocity direction, will indeed be small.

To trace the effect of image-point position error on the range error, it is only necessary to express the angle differences in the last two terms on the right-hand side of (15) as functions of the FOE and the image-point coordinates. Then the quantities $e_{v\theta}$, $e_{\theta\theta'}$ can be calculated in terms of derivatives of those functions.

If $\underset{\sim}{\rho}$ is the two-dimensional radius vector from the origin of the image-plane coordinate system to an image point in the first frame and $\underset{\sim}{\rho}'$ is that from the origin of the translated image-plane coordinate system to an image point in the second frame, then according to (4), normalized by setting D equal to 1, one has

$$\mathbf{R}/z = \underset{\sim}{\rho} + \mathbf{k}$$

and (16)

$$\mathbf{R}'/z' = \underset{\sim}{\rho}' + \mathbf{k} \quad .$$

Equations (16) determine the relation between the image-point positions and the angle difference $\theta - \theta'$ appearing in (15).

The cross product of the left- and right-hand sides of the first equation in (16) with the corresponding left- and right-hand sides of the second equation leads to

$$\mathbf{R} \times \mathbf{R}'/zz' = \underset{\sim}{\rho} \times \underset{\sim}{\rho}' + \underset{\sim}{\rho} \times \mathbf{k} + \mathbf{k} \times \underset{\sim}{\rho}' \, . \tag{17}$$

Taking the magnitude of the vector expression on each side of (17), keeping in mind that $\underset{\sim}{\rho}$ and $\underset{\sim}{\rho}'$ are both orthogonal to $\mathbf{k}$, gives the desired relation in terms of the image-point polar coordinates (ρ, θ_i) and (ρ', θ_i'):

$$|\sin(\theta - \theta')| = \frac{zz'}{RR'}\sqrt{\rho^2\rho'^2 \sin^2(\Delta\theta_i) + \rho^2 + \rho'^2 - 2\rho\rho' \cos(\Delta\theta_i)} \, , \tag{18}$$

where $\Delta\theta_i$ has been written for the angle difference $\theta_i - \theta_i'$.

From the fact that the two-dimensional image plane radius vector $\underset{\sim}{\rho}_F$ to the FOE is given by

$$\underset{\sim}{\rho}_F + \mathbf{k} = \mathbf{v}/v_z \, , \tag{19}$$

a similar relation can be derived for the angle difference $\theta_v - \theta$ in terms of the polar coordinates (ρ, θ_i) for image point and (ρ_F, θ_F) for the FOE. That is, taking the cross

product of both sides of the first equation in (16) with the corresponding sides of (19) leads to

$$\mathbf{v} \times \mathbf{R}/zv_z = \underset{\sim}{\rho}_F \times \underset{\sim}{\rho} + \underset{\sim}{\rho}_F \times \mathbf{k} + \mathbf{k} \times \underset{\sim}{\rho}_F \ ,$$

the magnitudes of both sides of which give the relation

$$|\sin(\theta_v - \theta)| = \frac{zv_z}{vR}\sqrt{\rho^2\rho_F^2 \sin^2(\Delta\theta_F) + \rho^2 + \rho_F^2 - 2\rho\rho_F \cos(\Delta\theta_F)} \ , \qquad (20)$$

where $\Delta\theta_F$ has been written for the angle difference $\theta_F - \theta_i$.

It is useful to note that the following substitutions transform (18) into (20):

v_z for z';

v for R';

ρ_F for ρ';

$\Delta\theta_F$ for $\Delta\theta_i$.

It is also useful to note that neither (18) nor (20) changes if the radial coordinates in either are interchanged. Thus an error analysis performed on

$$|\sin(\theta_1 - \theta_2)| = C\sqrt{\rho_1^2\rho_2^2 \sin^2(\Delta\theta) + \rho_1^2 + \rho_2^2 - 2\rho_1\rho_2 \cos(\Delta\theta)} \ , \qquad (21)$$

where C is a constant that has no effect on the fractional error contribution, will do for both (18) or (20). Also, the result obtained for one radial coordinate ρ_1 will do for the other radial coordinate, and the result obtained for one angle difference $\Delta\theta$ will do for the other angle difference.

In accordance with the general principle stated earlier, a logarithmic derivative of the right-hand side of (18) or (20) with respect to an image-point parameter, multiplied by the parameter error, gives the contribution of the parameter error to the fractional error of the quantity on the left-hand side. On the other hand, it is clear from (14) that the logarithmic derivatives of the left-hand sides of (18) and (20) contribute in the same way to the fractional error in the estimated range.

Thus, there are only two cases: the radial contribution $F_\rho(\rho_1, \rho_2, \Delta\theta)$ given by the logarithmic derivative of (21) with respect to ρ_1 and the angle difference contribution $F_\theta(\rho_1, \rho_2, \Delta\theta)$ given by the logarithmic derivative of (21) with respect to $\Delta\theta$. These quantities are

$$F_\rho(\rho_1, \rho_2, \Delta\theta) = \{\rho_1[\rho_2^2 \sin^2(\Delta\theta) + 1] - \rho_2 \cos(\Delta\theta)\}/\mathbf{D}$$

and (22)

$$F_\theta(\rho_1, \rho_2, \Delta\theta) = [\rho_1^2\rho_2^2 \sin(\Delta\theta)\cos(\Delta\theta) + \rho_1\rho_2\sin(\Delta\theta)]/\mathbf{D},$$

where

$$\mathbf{D} = \rho_1^2\rho_2^2 \sin^2(\Delta\theta) + \rho_1^2 + \rho_2^2 - 2\rho_1\rho_2\cos(\Delta\theta) \quad .$$

Note that the fractional range error given by (15) can be written

$$\varepsilon = \varepsilon_v/v + \varepsilon(\rho)\,[F_\rho(\rho, \rho_F, \theta_F - \theta_i) - F_\rho(\rho, \rho', \theta_i - \theta_i')] + \varepsilon(\rho')\,F_\rho(\rho', \rho, \theta_i - \theta_i')$$

$$+ \varepsilon(\rho_F)\,F_\rho(\rho_F, \rho, \theta_F - \theta_i) - \varepsilon(\theta_i)\,[F_\theta(\rho, \rho', \theta_i - \theta_i') + F_\theta(\rho, \rho_F, \theta_F - \theta_i)] \qquad (23)$$

$$+ \varepsilon(\theta_i')F_\theta(\rho, \rho', \theta_i - \theta_i') + \varepsilon(\theta_F)\,F_\theta(\rho, \rho_F, \theta_F - \theta_i)$$

in terms of absolute errors $\varepsilon(\rho)$, $\varepsilon(\rho')$, $\varepsilon(\rho_F)$, $\varepsilon(\theta_i)$, $\varepsilon(\theta_i')$, $\varepsilon(\theta_F)$ in the image-point and FOE radial and angular coordinates.

The error factors $F_\rho(\rho_1, \rho_2, \Delta\theta)$ and $F_\theta(\rho_1, \rho_2, \Delta\theta)$ will both be greatly magnified for parameter values that make the denominator $\mathbf{D}$ in both equations of (22) small, unless they also cause the numerators to decrease at the same rate. On setting $\mathbf{D}$ equal to zero, it will be found that only when the radial coordinates ρ_1 and ρ_2 are equal will a real angle solution of the resulting equation for $\Delta\theta$ exist and that the only solution in that case is $\Delta\theta = 0$.

In the limit of small $\Delta\theta$, dropping powers of $\Delta\theta$ higher than the first, (22) becomes

$$F_\rho(\rho_1, \rho_2, \Delta\theta) \sim 1/(\rho_1 - \rho_2)\,,$$
$$F_\theta(\rho_1, \rho_2, \Delta\theta) \sim \Delta\theta\rho_1\rho_2(1 + \rho_1\rho_2)/(\rho_1 - \rho_2)^2. \qquad (24)$$

It is evident from (24) that the case $\rho_1 = \rho_2$ is singular for both angular and radial image coordinate errors.

However, the angular coordinate error contribution shrinks if the angle difference or either radial coordinate decreases but the radial coordinate error contribution does not. Thus, for some geometric configurations radial image-point and FOE coordinate errors are more dangerous than angular coordinate errors. This is a particularly true when the radial coordinate values are small, which is always the case for the image points when the field of view is narrow.

The angular errors in (23) are errors in the difference between two angle coordinates. Therefore, errors in the angle coordinates of such a pair will tend to cancel out if they are in the same direction and have nearly the same magnitude. This may very well be the case for some errors due to fixed imperfections in the sensor's optical system.

Radial distortion due to an imperfect lens design will cause errors only in the radial image-point coordinates. Tangential and other less symmetric types of distortion, which occur for various different reasons mentioned in Section IV, can cause errors in both the radial and the angular image-point coordinates.

Since both ρ and ρ' will be small for a narrow field of view, (24) does not support the Ref. 2 claim that an optimal sensor velocity direction is one nearly parallel to the optical axis because that condition implies a small value for ρ_F. On the other hand, (24) does tend to support the Ref. 1 recommendation that the velocity vector be orthogonal to the optical axis, which implies a large value for ρ_F.

VI. RANGING ACCURACY

This section examines in more detail the effect of the error propagation estimates obtained in Section V on the accuracy of target ranging by means of passive image sensing. The discussion covers some general relations between the error and various geometric parameters that determine its magnitude. The results applied to numerical examples for two typical cases illustrate the dependence of ranging accuracy on geometry and the limitations thereby imposed in a particular application.

It is assumed that the sensor platform at an altitude h moves with a constant velocity v parallel to the ground. It is also assumed that the optical axis lies in the plane of the velocity vector and the vertical, that the object point for which the range is required is on the ground, and that the object point too lies in the same plane.

For the sake of completeness, Appendix B discusses the more general case in which the last two assumptions are dropped. However, no significant difference will be found between the general case and the more specialized case considered in this section, and the results obtained here should approximate, well enough to be regarded as entirely representative, most situations likely to be encountered in practice.

Figure 3 illustrates the geometrical configuration. It shows the X-Z plane of the optical coordinate system at two sensor positions separated by a time interval t, with the Z axis and the optical axis assumed to be coincident as usual.

The object point observed from the first sensor position is at a distance R_1, to be estimated by photogrammetric or optical flow analysis, in a direction given by the angle θ_1 relative to the Z, or optical, axis. As indicated in the figure, the line connecting the sensor and the object point also makes an angle ψ_1 with the horizontal. The same object point viewed by the sensor in its second position makes an angle θ_2 with the new position of the optical axis, and the line connecting the sensor in the second position with the object point makes an angle ψ_2 with the horizontal. The lines connecting the sensor in the two different positions with the object point meet in an angle α.

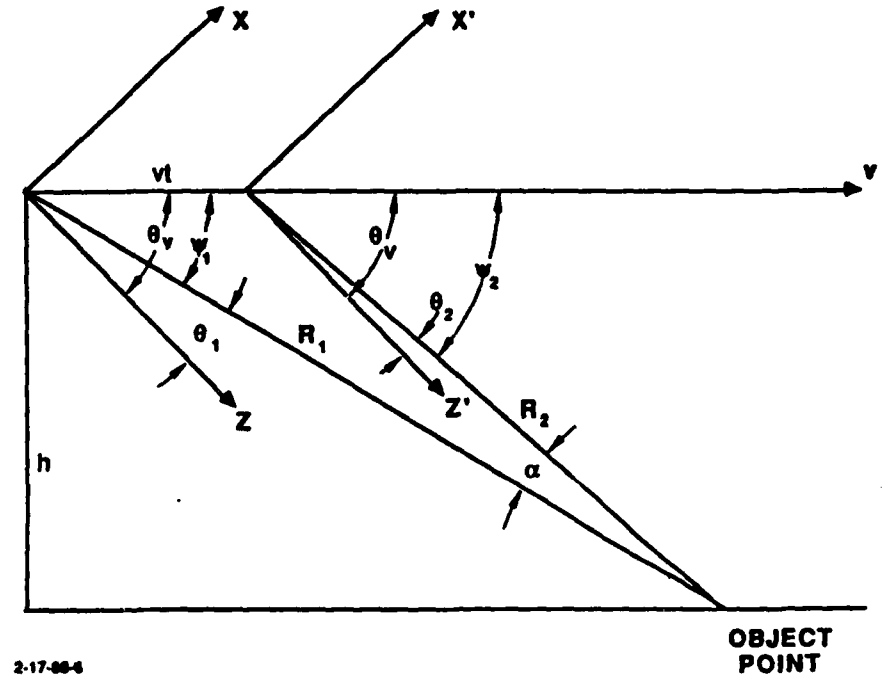


Figure 3. Image Angle Change During a Translation.

It is evident that when the sensor is in the first and second positions the radial coordinates ρ_1, ρ_2 of the image points corresponding to the object point are given by

$$\rho_1 = \tan \theta_1 \ ,$$

$$\rho_2 = \tan \theta_2 \ . \qquad (25)$$

It is also evident that

$$\theta_1 + \psi_1 = \theta_2 + \psi_2 \ ,$$

from which it follows that

$$\theta_2 = \theta_1 - \alpha \ .$$

Therefore, the Eqs. (25) become

$$\rho_1 = \tan \theta_1 \ ,$$

$$\rho_2 = \tan (\theta_1 - \alpha) \ . \qquad (26)$$

That is,

$$\rho_2 = (\rho_1 - \tan \alpha)/(1 + \rho_1 \tan \alpha) \ . \qquad (27)$$

According to (24) in Section V, the fractional error propagation factor is given by the reciprocal of $\rho_1 - \rho_2$. Therefore, the value of this quantity is a primary interest. From (27) it follows that

$$\rho_1 - \rho_2 = [(\rho_1^2 + 1)/(1 + \rho_1 \tan \alpha)] \tan \alpha \ . \tag{28}$$

For a sensor with a narrow field of view $\rho_1 << 1$. If α is acute and not near 90 deg, then (28) implies that

$$\rho_1 - \rho_2 \sim \tan \alpha \ . \tag{29}$$

It can be seen from Fig. 3 that

$$R_1 \sin \alpha = vt \sin (\psi_1 + \alpha) \ ,$$

from which it follows that

$$\tan \alpha = vth \Big/ \left(R_1^2 - vt \sqrt{R_1^2 - h^2} \right) \ . \tag{30}$$

It is evident from (24) and (29) that (30) expresses a fundamental relationship between the error and the important configuration parameters.

From its derivative with respect to t in (30), it can be seen that tan α increases, and therefore that the error decreases, monotonically as the time interval t between observations increases. In fact, tan α becomes infinite, thereby reducing the error factor to zero, when

$$t = R_1^2 \Big/ \left(v \sqrt{R_1^2 - h^2} \right) \ .$$

However, it is easy to see from Fig. 3 that the object-point image will be observable over this time interval only if the sensor field of view is at least 90 deg. In this event, since the corresponding value of ρ_1 would be at least 1, the first-case approximation of (29) would not be valid, but the exact Eq. (28) would give $\rho_1 + 1/\rho_1$ for the error factor.

Thus, the error factor is limited by the interval t between observation times, and t is limited either by the sensor field of view or by the requirement that the velocity remain constant during the interval. Assuming that the velocity can be held constant, Fig. 3 shows that the maximum possible value of α is the field-of-view angle φ_{FOV}, in terms of which the required time interval t_{MAX} would be given by

$$t_{MAX} = R_1^2 \tan \varphi_{FOV} \Big/ \left[v \left(h + \tan \varphi_{FOV} \sqrt{R_1^2 - h^2} \right) \right] \ . \tag{31}$$

By differentiating the right-hand side of (30) with respect to R_1 and setting the result equal to zero, the object-point range at which the value of tan α is maximized, and therefore the error is minimized, can be found in terms of a specified time interval between observations and the platform velocity and altitude. If $vt < 2h$, the result is given by

$$R_{MAX} = \sqrt{h^2 + (vt/2)^2} \; ; \tag{32}$$

otherwise, for two values of R_1^2,

$$R_1^2 = vt\left[vt \pm \sqrt{(vt)^2 - 4h^2}\right]/2 \; ,$$

tan α is infinite. The value of tan α corresponding to (32) is given by

$$\tan \alpha_{MAX} = vth/[h^2 - (vt/2)^2] \; , \tag{33}$$

which is a maximum if $vt < 2h$.

Actually, it is obvious from the geometry that, independently of the range, the time interval t must not be less than a certain duration for a given value of tan α. To find this minimum value of t, the first step is to solve (30) for R_1^2, obtaining

$$R_1^2 = vt\left\{\left[2h \cot \alpha + vt \pm \sqrt{(vt)^2 + 4vth \cot \alpha - 4h^2}\right]/2\right\} \; . \tag{34}$$

To get a real value for R_1 it is clear from (34) that the condition

$$(vt)^2 + 4vth \cot \alpha - 4h^2 > 0 \tag{35}$$

must be satisfied. A little manipulation shows that (35) is equivalent to

$$t \geq (2h/v) \tan \alpha/2 \sim (h/v) \tan \alpha \; . \tag{36}$$

For the minimum time interval given by (36), the range is given by

$$R_1 \sim (h \sec \alpha)/2, \tag{37}$$

which implies the geometrically obvious fact that for the image point to remain within the field of view, it must lie in a nearly vertical direction relative to the platform velocity.

The argument leading to (31) implies that the maximum value of tan α, and therefore the minimum error, occurs when the time duration t is such that the image moves exactly across the field of view. This fact, combined with the implication of (37), clearly supports the standard photogrammetric procedure of looking down rather than forward to obtain mapping accuracy, since that procedure leads to the shortest required time interval between frames for a given error factor.

It is instructive to use (24) and (30), along with (23), in some numerical examples that illustrate how sensitive the range error due to optical image distortion alone may be to typical geometrical and physical parameters. Tables 1 and 2 present the limits on radial distortion that would be necessary to obtain 90 percent accuracy in a range estimate based on data observed from two different sensor platforms. One, moving with slow speed at a low altitude, is consistent with a helicopter, and the other, moving faster at a higher altitude, is consistent with a light aircraft. The tables indicate clearly how difficult it should be to obtain accurate range estimates using a sensor with a narrow field of view when the data provided by the sensor is based on images viewed within a short time interval.

The results shown in the tables are given for various angular positions of the object point relative to the optical axis when observed in the first of two frames. The results are also given for the two cases in which the time interval between the two frames is either 1 second or 2 seconds.

Table 1. Maximum Image Distortion (mrad) for a 10 Percent Range Error; Platform Altitude 30 m, Velocity 50 knots

Initial Image Angle Off Optical Axis, deg	Time, sec	Maximum Image Distortion, mrad		
		500 m Range	1,000 m Range	1,500 m Range
2	1	0.326	0.079	0.035
	2	0.689	0.163	0.071
6	1	0.239	0.080	0.035
	2	0.696	0.165	0.072
12	1	0.340	0.083	0.036
	2	0.719	0.170	0.074

Table 2. Maximum Image Distortion (mrad) for a 10 Percent Range Error; Platform Altitude 500 m, Velocity 200 knots

Initial Image Angle Off Optical Axis, deg	Time, sec	Maximum Image Distortion, mrad		
		3,000 m Range	5,000 m Range	10,000 m Range
2	1	0.593	0.210	0.052
	2	1.228	0.430	0.105
6	1	0.598	0.212	0.053
	2	1.239	0.434	0.106
12	1	0.618	0.220	0.054
	2	1.279	0.448	0.110

VII. SUMMARY AND CONCLUSIONS

With two-dimensional image data obtained passively at two different sensor locations and certain other information, it is possible to locate the position of an object point fixed in three-dimensional space relative to either sensor position. The required additional information is equivalent to knowledge of the transformation between fixed optical coordinate systems at the two sensor positions.

Six independent parameters characterize the transformation: three components of a linear displacement plus a rotation through some angle about an axis which two other angles determine. Image data alone, observed at the two sensor locations, can yield five of the six transformation parameters, but not the remaining one.

When the two sensor positions are those of a single moving sensor at the beginning and the end of a known, sufficiently short time interval, a calculation based on the concept of optical flow theory can, in principle, give the five obtainable transformation parameters. In general, a system of at least five quadratically coupled algebraic equations, which can be set up if data from at least five image points and their optical flow displacements are available, will determine those quantities.

It may also be possible to calculate them as part of the solution of a system of 11 linear algebraic equations, which can be set up with similar data from at least 11 image points. However, if data from at least nine image points are available, the method derived in Appendix A will provide a closed-form solution for the five transformation parameters.

A common method used in photogrammetry to obtain map scale can determine the remaining transformation parameter, but it requires knowledge of the actual locations of three known control points imaged in a single stationary scene observed by a sensor. Otherwise, the procedure based on optical flow can determine the direction but not the magnitude of the moving sensor's velocity.

A number of error sources are associated with the optical characteristics of any imaging sensor and contribute to the error in estimating the position of an object in space: its range, in particular. To estimate range with any accuracy, it is necessary to perform

certain preflight and possibly in-flight calibrations of the optical equipment to account for some of the distortion introduced into the image data. Algorithms for calculating the range must take into account the information acquired by this means.

In general, for best results the displacement between the two sensor positions should be perpendicular to the optical axis after the effect of any rotation is removed from the image data. That is, the sensor motion should be such that the FOE lies well outside the field of view.

When the two images used to estimate the range of an object point are not far apart, the range error due to radial distortion in the optical system will be magnified by a correspondingly large factor, which, in fact, becomes infinite as the difference between their radial coordinates approaches zero. On the other hand, errors due to tangential distortion may cancel if the effect of that distortion is the same for both images.

Equations (22) and (23) in Section V give an estimate of the error in the range calculated by means of optical flow theory. The estimate depends on errors in the image-plane coordinates of image points and the FOE. Normal optical system imperfections, listed in Section IV, are an important but often neglected source of such errors.

REFERENCES

1. C.C. Slama, ed., *Manual of Photogrammetry*, 4th ed., American Society of Photogrammetry, Falls Church, Virginia, 1980.

2. M.E. Bazkos et al., *Passive Ranging in Outdoor Environment*, Honeywell Systems and Research Center, Signal and Image Processing, 2600 Ridgeway Parkway, N.E., MN 17-2357, Minneapolis, Minnesota 55413 (unpublished).

3. R.M. Harelick, "Using Perspective Transformations in Scene Analysis," *Computer Graphics and Image Processing*, 13, 191-221 (1980).

4. K. Prazdny, "Determining the Instantaneous Direction of Motion from Optical Flow Generated by a Curvilinearly Moving Observer," *Techniques and Applications of Image Understanding*, Society of Photo-Optical Instrumentation Engineers, 281, 199-206 (1981).

5. R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Vol. I, English ed., Interscience, New York, 1953.

APPENDIX A

CALCULATION OF THE VECTORS v AND ω

CALCULATION OF THE VECTORS v AND ω

This appendix derives what, from several points of view, is probably the most effective method of calculating the translation and rotation vectors, the components of which satisfy (9). In terms of all of the available data, the method leads to a closed-form solution and is therefore free of any convergence problems or unanticipated numerical instabilities.

It is assumed initially that the optical flow data consists of the values of the displacements dx and dy measured at 10 or more image points. The first step is to represent dx and dy as functions of the image-point coordinates x, y by third-degree polynomials

$$dx = A_1 + B_1x + C_1y + D_1x^2 + E_1xy + F_1y^2 + G_1x^3 + H_1x^2y + J_1xy^2 + K_1y^3,$$

$$dy = A_2 + B_2x + C_2y + D_2x^2 + E_2xy + F_2y^2 + G_2x^3 + H_2x^2y + J_2xy^2 + K_2y^3 \quad . \tag{A-1}$$

If the number of image-point measurements is greater than 10, higher degree approximations, or perhaps a least square error fit, can be used to obtain the representation; otherwise, any type of polynomial interpolation (e.g., Lagrange, spline) that suits the image contrast variation can be used.

Once the polynomial fit to the measured data is accomplished, the task of calculating the translation and rotation vectors is reduced to solving systems of at most three linear algebraic equations at a time. Therefore, following the procedure described at length further on in this appendix, the remaining computation can be done quickly and as accurately as desired.

From the displacement definitions given by (7), with D set equal to 1, and the Weierstrass polynomial approximation theorem (cf. Ref. A-1, p. 65): if z is a continuous

function of x and y (it may be assumed for physical reasons that z is never zero) in the closed clear-aperture region of the image plane, a uniformly convergent approximation by polynomials is possible. If z is constant over the image plane, the displacements will actually be polynomials, but quadratic rather than cubic. This is a special case which will be treated later.

After multiplying through by the denominators on both sides, with some algebraic manipulation and rearrangement of terms (9) can be written

$$[(v_y - v_z y)\,dx + (v_z x - v_x)\,dy] + (v_y\omega_y + v_x\omega_x) - (v_x\omega_z + v_z\omega_x)\,x$$

$$- (v_y\omega_z + v_z\omega_y)\,y + (v_y\omega_y + v_z\omega_z)\,x^2 - (v_y\omega_x + v_x\omega_y)\,xy + (v_z\omega_z + v_x + v_x\omega_x)\,y^2 = 0. \qquad \text{(A-2)}$$

Note that the collection of terms not enclosed in brackets is a second-degree polynomial in x and y, the cubic terms that result from clearing fractions in (9) having canceled out.

On the other hand, after substitution from (A-1) for dx and dy, the collection inside the brackets is a fourth-degree polynomial in x and y. In fact, the fourth-degree terms in that polynomial together form a homogeneous polynomial P_4 given by

$$P_4 = G_2x^4 + (H_2 - G_1)\,x^3y + (J_2 - H_1)\,x^2y^2 + (K_2 - J_1)\,xy^3 - K_1y^4. \qquad \text{(A-3)}$$

If the polynomial representations of dx and dy were exact instead of approximate, the coefficient of each monomial term in (A-2) [including each of those in (A-3), of course] would be zero. Thus, from (A-3) it would follow that

$$G_2 = K_1 = 0,\ H_2 = G_1,\ J_2 = H_1,\ K_2 = J_1\ . \qquad \text{(A-4)}$$

Then (A-1) could be replaced by

$$dx = A_1 + B_1x + C_1y + D_1x^2 + E_1xy + F_1y^2 + G_1x^3 + H_1x^2y + J_1xy^2\ ,$$

$$dy = A_2 + B_2x + C_2y + D_2x^2 + E_2xy + F_2y^2 + G_1x^2y + H_1xy^2 + J_1y^3\ . \qquad \text{(A-5)}$$

In some sense the extent to which the relations (A-4) are satisfied provides a measure of the accuracy of the data in combination with that of the polynomial approximation. However, it might be better to consider an alternative polynomial approximation.

Instead of using data from a minimum of 10 image points to obtain a representation of dx and dy in the form (A-1), an alternative would be to use data from a minimum of 8 image points to obtain a representation in the form (A-5). This is equivalent to imposing a self-consistency boundary condition on the approximation.

After determining the explicit polynomial coefficients in (A-1) or (A-5), the next step is to combine terms within the brackets in (A-2) with like terms inside and outside the brackets and, as before, to set the resulting coefficient of each separate monomial equal to zero. The cubic terms lead to a system of four linear homogenous algebraic equations for the components of v:

$$\begin{aligned} G_2 v_x - G_1 v_y - D_2 v_z &= 0, \\ K_2 v_x - K_1 v_y + F_1 v_z &= 0, \\ H_2 v_x - H_1 v_y + (D_1 - E_2) v_z &= 0, \\ J_2 v_x - J_1 v_y + (E_1 - F_2) v_z &= 0. \end{aligned} \tag{A-6}$$

If the polynomial representations for dx and dy have the forms (A-5), according to (A-4) the system of equations will reduce to

$$\begin{aligned} G_1 v_y + D_2 v_z &= 0, \\ J_1 v_x + F_1 v_z &= 0, \\ G_1 v_x - H_1 v_y + (D_1 - E_2) v_z &= 0, \\ H_1 v_x - J_1 v_y + (E_1 - F_2) v_z &= 0. \end{aligned} \tag{A-7}$$

In either case, the result is a system of four linear homogenous algebraic equations in three unknowns, of the form

$$\begin{aligned} M_{11}v_1 + M_{12}v_2 + M_{13}v_3 &= 0, \\ M_{21}v_1 + M_{22}v_2 + M_{23}v_3 &= 0, \\ M_{31}v_1 + M_{32}v_2 + M_{22}v_3 &= 0, \\ M_{41}v_1 + M_{42}v_2 + M_{43}v_3 &= 0, \end{aligned} \tag{A-8}$$

wherein the v_1 represent components of v arranged in an arbitrary order, but in the same order in every equation. Assuming that v_3 is a nonzero component of v, if each equation of (A-8) is divided by v_3, the result is a system of four equations in two unknowns:

$$\begin{aligned} M_{11}w_1 + M_{12}w_2 &= -M_{13}, \\ M_{21}w_1 + M_{22}w_2 &= -M_{23}, \\ M_{31}w_1 + M_{32}w_2 &= -M_{33}, \\ M_{41}w_1 + M_{42}w_2 &= -M_{43}, \end{aligned} \tag{A-9}$$

where $w_1 = v_1/v_3$ and $w_2 = v_2/v_3$.

The linear system of equations (A-9) is overdetermined, but its least square error solution can be found by multiplying both sides by the transpose of the coefficient matrix on the left-hand side and solving the resulting system

$$\begin{aligned} Aw_1 + Bw_2 &= C, \\ Bw_1 + Dw_2 &= E, \end{aligned} \tag{A-10}$$

where

$A = M_{11}^2 + M_{21}^2 + M_{31}^2 + M_{41}^2$, $B = M_{11}M_{12} + M_{21}M_{22} + M_{31}M_{32} + M_{41}M_{42}$,

$D = M_{12}^2 + M_{22}^2 + M_{32}^2 + M_{42}^2$, $C = -M_{11}M_{13} - M_{21}M_{23} - M_{31}M_{33} - M_{41}M_{43}$,

$E = -M_{12}M_{13} - M_{22}M_{23} - M_{32}M_{33} - M_{42}M_{43}$.

As many as three different possible systems of the form (A-9) can be obtained from (A-7) or (A-8), depending on which component of v is chosen for v_3 (assuming that more than one differs from zero). For greatest accuracy, the choice should be that for which the corresponding coefficient determinant $AD-B^2$, which is guaranteed to be nonnegative, has the largest value.

To find the components of the rotation vector ω, it is only necessary to repeat the steps of combining like monomial terms in (A-2), this time concentrating on those of degree less than three, and once more setting the resulting coefficients equal to zero. This will provide a system of six linear algebraic equations in the three unknowns ω_x, ω_y, ω_z, since the components of v, having been calculated in the previous step, will be known parameters,

$$
\begin{aligned}
v_x\omega_x + v_y\omega_y &= A_2v_x - A_1v_y, \\
v_z\omega_y + v_y\omega_z &= -C_2v_x + C_1v_y - A_1v_z, \\
v_z\omega_x + v_x\omega_z &= -B_2v_x + B_1v_y + A_2v_z, \\
v_y\omega_y + v_z\omega_z &= D_2v_x - D_1v_y - B_2v_z, \\
v_x\omega_x + v_z\omega_z &= F_2v_x - F_1v_y + C_1v_z, \\
v_y\omega_x + v_x\omega_y &= -E_2v_x + E_1v_y + (C_2 - B_1)\,v_z.
\end{aligned}
\qquad \text{(A-11)}
$$

To find the least square error solution to (A-11), again the procedure is to multiply both sides by the transpose of the coefficient matrix. In this case, the system reduces to three equations for the three ω components:

$$\Omega_{11}\omega_x + \Omega_{12}\omega_y + \Omega_{13}\omega_z = \Gamma_1 ,$$
$$\Omega_{21}\omega_x + \Omega_{22}\omega_y + \Omega_{23}\omega_z = \Gamma_2 , \qquad \text{(A-12)}$$
$$\Omega_{31}\omega_x + \Omega_{32}\omega_y + \Omega_{33}\omega_z = \Gamma_3 ,$$

where

$\Omega_{11} = 2v_x^2 + v_y^2 + v_z^2$, $\Omega_{12} = 2v_xv_y$, $\Omega_{13} = 2v_xv_z$,

$\Omega_{21} = 2v_xv_y$, $\Omega_{22} = v_x^2 + 2v_y^2 + v_z^2$, $\Omega_{23} = 2v_yv_z$,

$\Omega_{31} = 2v_xv_z$, $\Omega_{32} = 2v_yv_z$, $\Omega_{33} = v_x^2 + v_y^2 + 2v_z^2$,

and

$$\Gamma_1 = (A_2 + F_2)\, v_x^2 - (A_1 + F_2 + F_1)\, v_xv_y + (C_1 - B_2)\, v_xv_z + E_1v_y^2 + C_2v_yv_z + A_2v_z^2,$$
$$\Gamma_2 = -E_2v_x^2 + (A_2 + D_2 + E_1)\, v_xv_y - B_1\, v_xv_z - (A_1 + D_1)\, v_y^2 + (C_1 - B_2)\, v_yv_z - A_1v_z^2,$$
$$\Gamma_3 = -B_2v_x^2 + (B_1 - C_2)\, v_xv_y + (A_2 + D_2 + F_2)\, v_xv_z + C_1v_y^2 - (A_1 + D_1 + F_1)\, v_yv_z + (C_1 - B_2)\, v_z^2 .$$

The solution of the system of equations (A-12) is straightforward because its coefficient matrix is nonsingular. Demonstrating this is not difficult.

The statement that the coefficient matrix of (A-12) is nonsingular is equivalent to the statement that it has rank 3. That this matrix has rank 3 follows from the fact that it must have the same rank as the coefficient matrix of (A-11), which, it can be shown, has rank 3.

To establish the last fact it is sufficient to find a nonsingular 3 by 3 submatrix. For this purpose just two cases need to be considered: the submatrix consisting of the first three rows of the coefficient matrix of (A-11) is singular only if, first, v_z vanishes or, second, v_x and v_y both vanish. Since, by hypothesis, translational motion exists, these cases are mutually exclusive.

In the first case, if v_x does not also vanish, the submatrix consisting of the third, fifth, and sixth rows is nonsingular, or, if it is v_y that does not also vanish, the submatrix consisting of the second, fourth, and sixth rows is nonsingular. In the second case, the submatrix consisting of the second, third, and fourth rows is nonsingular.

For the special case mentioned earlier in which Z is constant for all image points, dx and dy are both second-degree polynomials in x and y. Then in (A-1) the coefficients $G_{1,2}$, $H_{1,2}$, $J_{1,2}$, $K_{1,2}$ are all zero, as are the coefficients of all monomials of higher degree in the polynomial approximations to dx and dy.

The coordinate system transformation vectors are then given by

$$\begin{aligned} \omega_x &= E_1 \ , \\ \omega_y &= -D_1 \ , \\ \omega_z &= -B_2 \ , \\ -v_x/Z &= A_1 - D_1 \ , \\ -v_y/Z &= A_2 - E_1 \ , \\ -v_z/Z &= -B_1 \ . \end{aligned} \tag{A-13}$$

It should be noted that the last three equations of (A-13) determine the direction but not the magnitude of v, since Z, although a constant, is still unknown.

REFERENCE

A-1. R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Vol. I, English ed., Interscience, New York, 1953.

APPENDIX B

RANGING ACCURACY IN A GENERAL GEOMETRICAL CONFIGURATION

RANGING ACCURACY IN A GENERAL GEOMETRICAL CONFIGURATION

This appendix extends the analysis of Section VI to the most general case in which the sensor platform velocity is constant relative to an object point during the time interval between two observations of the point. That is, it is no longer assumed that the passive ranging target lies in the plane of the velocity vector and the optical axis of the sensor.

Figure 3, slightly reinterpreted, will do for illustrating the more general geometrical configuration as well as the more restrictive one treated in Section VI. The only significant change is that the object point need not be in the plane determined by the velocity vector v and the optical, or Z, axis.

The velocity vector is no longer assumed parallel to the ground; therefore, the line labeled h in the figure, although assumed to be orthogonal to v, is not necessarily vertical. In fact, the ground line on which it was originally assumed that the object point was located is irrelevant in the present discussion.

The line labeled h is defined by the following requirements. It lies in the plane determined by v and the optical axis, it is perpendicular to v, and its length is determined by a perpendicular dropped from the object point.

The optical coordinate system, which is Cartesian and right handed as usual, includes the optical axis as the Z axis. The X axis is defined by requiring that it lie in the plane determined by v and the Z axis, and that it make an acute angle with v. The Y axis is then defined by the requirement that the coordinate system be right handed, i.e., pointing out of the page in Fig. 3. The optical coordinate system can be regarded as the result of a rotation about the Y axis of a coordinate system for which the other two axes consist of v and the line labeled h.

For the discussion in Section VI, the relations

$$\theta_v = \theta_1 + \psi_1 ,$$

$$\theta_v = \theta_2 + \psi_2$$

were implicitly assumed. However, for the discussion in this appendix these relations are no longer valid because the three angles in each are in general the vertex angles of a pyramid.

Nevertheless, with this interpretation the relations

$$\rho_1 = \tan \theta_1,$$
$$\rho_2 = \tan \theta_2 \qquad \text{(B-1)}$$

are still valid. On the other hand, it is no longer necessarily true that $\rho_1 - \rho_2$ is approximately equal to $\tan \alpha$.

Therefore, investigating the behavior of the error factor as the time t between the two sensor observations increases requires a somewhat different approach from the one adopted in Section VI. For this purpose the relation

$$R_2 \cos \theta_2 = R_1 \cos \theta_1 - vt \cos \theta_v \qquad \text{(B-2)}$$

is useful.

The quantities with subscript 1 in (B-2), as well as θ_v, are independent of t. Therefore, differentiating both sides of (B-2) with respect to t leads to the equation

$$\dot{R}_2 \cos \theta_2 - R_2 \dot{\theta}_2 \sin \theta_2 = - v \cos \theta_v \quad . \qquad \text{(B-3)}$$

On the other hand, since the object point is fixed, it follows that

$$\dot{R}_2 = - v \cos \psi_2 \quad . \qquad \text{(B-4)}$$

Substituting from (B-4) into (B-3) then provides the equation

$$\dot{\theta}_2 = v(\cos \theta_v - \cos \psi_2 \cos \theta_2)/R_2 \sin \theta_2 \quad : \tag{B-5}$$

It follows from (B-5) that

$$\dot{\theta}_2 \leq O$$

if (B-6)

$$\cos \theta_v \leq \cos \psi_2 \cos \theta_2$$

since $\sin \theta_2$ is positive and can only vanish when the Z' axis, v, and the line of sight to the object point in the second observation position are coplanar, in which case the configuration must be the one treated in Section VI.

It follows from (B-1) that the error factor decreases as long as the inequality condition in (B-6) is satisfied, and it reaches a minimum when the equality condition is satisfied. That is, the ranging accuracy improves as the time interval between sensor observations increases from zero, but only for a while, and there is an optimum time interval for which the error sensitivity is a minimum. The major difference between this conclusion and a similar one reached in Section VI is that for an object point sufficiently far off track it is conceivable that the image point may still be within the sensor field of view when the optimum interval between observations actually occurs.